基于改进Linknet网络的黄土高原苹果园精准提取

张志博[1]，赵西宁[2*]，高晓东[2]，张 利[1]，杨孟豪[2]

（1. 西北农林科技大学 水利与建筑工程学院，陕西杨凌 712100；
2. 中国科学院水利部水土保持研究所，陕西杨凌 712100）

摘 要：黄土高原近20年来苹果栽植面积迅猛增加，对区域生态水文和社会经济发展均产生了重要影响。但该区域果园地块小且场景复杂，仅有县/市尺度统计数据，尚无苹果园实际的空间分布信息。为此，本研究建立了无人机低空遥感影像专业数据集。融合迁移学习与深度学习方法，将残差神经网络ResNet34网络迁移到Linknet网络，得到R_34_Linknet网络。将R_34_Linknet网络与5种常用的深度学习语义分割模型SegNet、FCN_8s、DeeplabV3+、UNet和Linknet应用于黄土高原苹果园空间分布提取，表现最好的模型为R_34_Linknet，其在测试集上的调和平均值F_1为87.1%，像素准确度PA为92.3%，均交并比MIoU为81.2%，频权交并比FWIoU为85.7%，平均像素准确度MPA为89.6%。将空间金字塔池化结构（Atrous Spatial Pyramid Pooling，ASPP）与R_34_Linknet网络相结合，扩大网络的感受野，得到R_34_Linknet_ASPP网络；然后对ASPP结构进行改进，得到R_34_Linknet_ASPP+网络。对比三种网络性能，表现最优的为R_34_Linknet_ASPP+，在测试集上F_1为86.3%，PA为94.7%，MIoU为82.7%，FWIoU为89.0%，MPA为92.3%。使用R_34_Linknet_ASPP+在长武县王东沟和白水县通积村提取苹果园面积精度分别为94.22%和95.66%。本研究提出的R_34_Linknet_ASPP+方法提取到的苹果园更加准确，苹果园地块边缘处效果更好，可作为黄土高原苹果园空间分布制图等研究的技术支撑和理论依据。

关键词：无人机遥感；苹果园提取；深度学习；黄土高原；迁移学习；残差神经网络；语义分割

中图分类号：S126；T18 **文献标志码：**A **文章编号：**SA202206001

引用格式：张志博，赵西宁，高晓东，张利，杨孟豪．基于改进Linknet网络的黄土高原苹果园精准提取[J]. 智慧农业(中英文)，2022, 4(3): 95-107.

ZHANG Zhibo, ZHAO Xining, GAO Xiaodong, ZHANG Li, YANG Menghao. Accurate extraction of apple orchard on the loess plateau based on improved linknet network[J]. Smart Agriculture, 2022, 4(3): 95-107. (in Chinese with English abstract)

1 引 言

退耕还林（草）工程实施20多年以来，黄土高原苹果产业发展迅猛，已经发展成为中国乃至全球最大的优质苹果生产基地。以2018年为例，黄土高原苹果栽培面积和产量分别为渤海湾产区的2.11倍和1.47倍[1]。苹果产业的迅猛发展一方面极大增加了农民收入，助力脱贫攻坚与乡村振兴，另一方面则不可避免地改变了区域生态水文过程，造成了负面的水循环效应[2]。通过合理配置苹果园的分布格局，实现苹果产业适水发展成为黄土高原苹果产业可持续发展的关键所在。但

收稿日期：2022-06-01

基金项目：国家重点研发计划项目（2021YFD1900700）；国家杰出青年科学基金项目（42125705）

作者简介：张志博（1995—），博士研究生，研究方向为农业遥感技术与图像分割。E-mail：zhang_zhibo@nwafu.edu.cn

*通信作者：赵西宁（1976—），博士，研究员，研究方向为农业水土资源利用与调控。E-mail：zxn@nwafu.edu.cn

是黄土高原地形复杂多变，苹果园空间格局破碎，找到一种快速准确获取区域苹果园空间分布的方法，是该区域苹果适水发展的重要前提。

传统黄土高原苹果园空间种植面积信息的获取主要依靠地方行政单位统计，各级上报或者按一定比例抽样调查来获得[3,4]。这些方法不但耗财、耗物，而且难以获取流域和区域果园种植准确面积。随着无人机遥感和卫星遥感平台的快速发展，传统机器学习方法（如支持向量机提取法、随机森林提取法、最大似然提取法）结合遥感技术[5-8]已被广泛用来提取苹果园空间分布信息，但是这些方法在提取精度差和效率方面还存在不足。目前，深度学习是人工智能领域的研究热点，卷积神经网络作为深度学习的重要分支，在图像分类[9,10]和语义分割[11,12]等领域取得了许多成果。分类网络后3层都是一维向量，计算方式采用全连接，因此丢失了二维信息，而分割网络将后3层全部转换为1×1的卷积核所对应同等向量长度的多通道卷积层，即把全连接层换成全卷积层。此外，采用全卷积神经网络（Fully Convolutional Networks，FCN）的语义分割是逐像素的分类，能够实现遥感影像的精准分割，相对于传统方法以及卷积神经网络，在植被提取上具有明显优势[13]。

Olaf等[14]在FCN的基础上修改并扩张了FCN网络得到UNet网络，使其在使用少量数据进行训练的情况下能够得到精确的分割结果[15]。编码网络和解码网络的对称结构构成了SegNet的主要部分，除此之外还有一些输出层[16]。FCN、UNet、SegNet网络的主干部分为超分辨率测试序列（Visual Geometry Group，VGG）中的VGG16[17]，但是当网络达到一定深度时，会造成退化问题，表现反而不如浅层网络，而将ResNet网络引入图像分割领域[18]，很好地解决了这种退化问题，提高了图像分类的提取精度[19]。为了解决特征图尺寸和感受野之间的矛盾，Deeplab系列引入了空间金字塔池化结构（Atrous Spatial Pyramid Pooling，ASPP）[20-22]。但Deeplab系列网络层数太多，而Chaurasia等[23]提出的LinkNet网络模型层数太少，并且LinkNet网络中没有ASPP结构，导致最终的分割效果也不好[24]。

目前，虽然卫星遥感是大面积农作物分类的主要手段，但存在成本昂贵、回访周期长和空间分辨率低等问题，且过于依赖卫星过境时的天气状况，实时性和准确性均受限。而无人机遥感具有空间分辨率高、周期短、灵活性高、受云层和天气影响小等优势，弥补了传统卫星遥感的不足，成为目前小区域农业遥感数据获取的主要手段。近年来，不需要人工选取特征参数、并且可以充分利用无人机超高分辨率的特点的深度学习语义分割算法的[25,26]出现和发展为超高分辨率无人机图像分割与分类提供了新的思路。因此，本研究针对黄土高原苹果园地块小且空间分布格局破碎[27,28]，采用无人机获取苹果园遥感影像，将RestNet34迁移到Linknet网络，构建R_34_Linknet网络，并结合ASPP用于黄土高原无人机影像苹果园的提取，通过试验证明方法的有效性，以为厘清黄土高原苹果园分布格局提供技术支撑。

2 材料与方法

2.1 试验区域与数据集构建

研究区域为黄土高原苹果主要栽植区。该区最冷月平均气温低于−5.0 °C、月平均气温高于10 °C不超5个月、年平均降水量小于470 mm[29,30]，具有典型的大陆季风气候特征。在地处黄土高原的12个市/县选择不同林龄、不同品种、不同管理方式的28个苹果园采集点作为研究对象，其分布如表1所示，采样时间为2020年5月下旬至2020年10月上旬。

飞行平台为大疆御2 Pro四旋翼无人机系统，三轴稳定云台，俯仰角度为−90°～+30°，其机身小巧、可折叠、易操作，自重仅为907 g。最大水平飞行速度72 km/h，最长飞行时间为31 min。

表 1　黄土高原不同市/县采样点数量分布

Table 1 Quantity distribution of sampling points in different counties on the Loess Plateau

县	通渭	正宁	扶风	长武	耀州	白水	洛川	延安	子州	绥德	米脂	灵宝
采样点数量/个	3	1	3	2	2	3	2	6	2	1	2	1

影像传感器为哈苏1英寸CMOS可见光RGB相机，有效像素2000万，视角约77°，机载内存64 GB。拍摄时，天气均晴朗，地面风速均小于2级，满足航摄要求。于飞行高度80～120 m，采集苹果园图像300幅，预处理之后图像像素为256×480，使用Python第三方开源软件Labelme进行人工标注。将图像中的像素分为苹果园和背景两类，ID值分别为0和1。采集的图像包含各种场景苹果园以及各种代表性干扰的图像，例如不同的光照强度、山的阴面和阳面、背景植物的干扰、果园长势不同和苹果有无套袋的变化等具有黄土高原苹果园多个特征的图像。

航拍时，因风速变化、光照变化以及无人机拍照过程中飞行姿态的调整都会影响影像的清晰度，本研究利用几何变换和饱和度变换扩充数据集，通过饱和度变换（0.1、0.5、1.5）、图像旋转、缩小填零和放大截取的方法将采集的图像扩充了2倍生成了600幅新图像，加上原始的300幅图像，数据集总共为900幅。

为了更好地利用训练集数据，本研究采用9折交叉验证法划分数据集。首先将全部图像随机分成9份，每份100幅图像，每份图像从原图和扩大后图像按照4:6比例挑选，每份图像包含各种场景的苹果园，含有间作果园、山地果园、不同果龄果园和复杂植物背景下果园等，然后不重复地每次取其中1份做测试集，其余8份做训练集，如此循环9次，保证所有的图像都进行过训练和测试，得到9个评价模型，最后将9次评价的性能均值作为最后评价结果。

2.2 苹果园整体提取方案

图1为本研究苹果园提取的整体方案，无人机影像经过添加ASPP的R_34_Linknet网络，ASPP位于R_34_Linket网络的编码器与解码器中间，ASPP在不引入额外参数的前提下可以增大网络感受野和提升边缘分割效果。训练时，将解码器输出的预测图与GT标签经损失函数进行对比，直至得到最小值，若不是最小值则进行反向传播调参，得到最终的预测图。

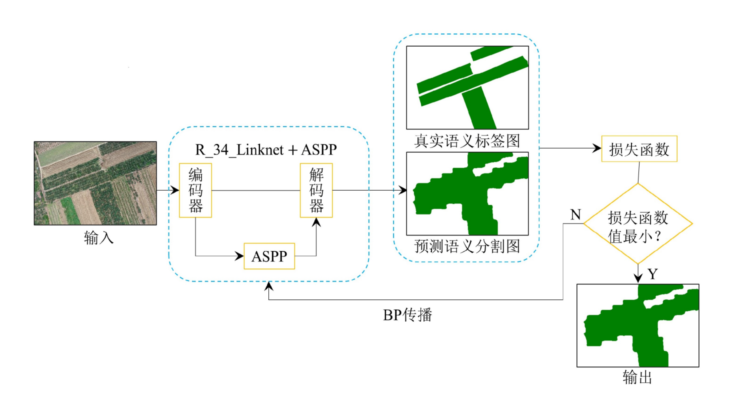


图 1　苹果园提取整体方案

Fig. 1 The overall extraction scheme of apple orchard

2.3 R_34_Linknet和ASPP

Linknet网络在U型全卷积神经网络的基础上引入了ResNet，实现了像素级的分类。原始的LinkNet使用了ResNet18作为编码器，能够在移动设备上实施运行的低功耗语义分割网络，但是

准确度不高且表征能力较弱。本研究分别把RestNet34、RestNet50、RestNet101、RestNet152迁移到LinkNet网络作为编码器，试验结果表明在本研究数据集上，RestNet34相较于RestNet50、RestNet101、ResNet152效果稍好，且ResNet34网络结构更简单、参数量更少，因此，本研究选择ResNet34作为Linknet的编码器，构成R_34_Linknet，提升网络的整体性能和精度。R_34_Linknet网络由两部分组成，第一部分是编码器，从一个初始块开始，对输入图像进行卷积，卷积核大小为7×7，步长为2，初始块后接一个3×3步长为2的最大池化层，后面的部分由4个编码层组成；第二部分是解码器，解码层由1层反卷积和2层卷积组成，共4个解码层。

在R_34_Linknet网络训练时，将$3 \times H \times W$的无人机图像通过第1个步长为2、卷积核尺寸为7×7、通道数为64的卷积层计算，输出像素尺寸变为$\frac{H}{2} \times \frac{W}{2}$，通道数变为64。计算方法如公式（1）。

$$X_j^t = f\left(\sum_{i=1}^{n} X_i^{t-1} \otimes E_{ij}^t + B_j^t\right) \quad (1)$$

其中，X_j^t为第t层输出的第j个特征图；f为激活函数$ReLU$；n为卷积核数量；X_i^{t-1}为第$t-1$层的第i个通道影像；E_{ij}^t为第t层卷积核；$\otimes$为卷积运算符号；B_j^t为第t层卷积核后第j个特征图的偏置。

为了增强模型的鲁棒性，减少参数的数量，防止过拟合现象的发生，在每个卷积层后面增加最大池化层。池化步长为2，尺寸为3×3。图像经池化操作之后，尺寸变为$\frac{H}{4} \times \frac{W}{4}$，通道数仍为64。

在编码器中，4个编码层输出特征图通道数分别为64、128、256、512，扩展到编码之前的8倍。在解码器中，4个解码层输入特征图通道数分别为512、256、128、64，特征图经过解码器后，尺寸和通道数变为与编码前相同。

解码器的输出图像进入反卷积层，可以还原特征图尺寸大小，同时节约计算成本。特征图经过解码操作之后图像尺寸还原为输入网络之前大小$H \times W$，通道数变为1，此时输出苹果园提取的预测图。

与传统的卷积算法相比，空洞卷积能够在不增加参数与计算量的情况下，扩大网络的感受野，更精确地定位目标，更好地捕捉多尺度上下文信息[31]。由空洞卷积获得的特征图与输入特征图尺寸相同，但每个输出的神经元拥有更大的感受野，因此可以获取更多细节信息同时降低分辨率的损失。空洞卷积可以设置不同空洞率，其具体操作为在卷积核中插入0，使卷积核尺寸扩大，因此设置不同空洞率时，网络感受野就不同，不同尺度上下文信息可以被更好地获取。

空洞卷积解决了感受野和特征图分辨率之间的矛盾，在卷积神经网络中，决定某一层输出结果中一个元素所对应的输入层的区域大小，被称为感受野。通常来说，大感受野的效果要比小感受野更好。输入图片的感受野约定为1，即$RF_1 = 1$，计算如公式（2）所示。

$$RF_{n+1} = RF_n + (kernel_size - 1) \times stride \quad (2)$$

其中，RF_{n+1}为第$n+1$层特征图感受野；RF_n为第n层特征图感受野；$kernel_size$为卷积核尺寸；$stride$为卷积核步长。

假设原始特征为feat0，卷积核尺寸为3×3，首先用空洞率为0的空洞卷积生成feat1，feat1上相对feat0的感受野为3×3；然后使用空洞率为2的卷积计算feat1生成feat2，让第一次空洞卷积的卷积核尺寸等于第二次空洞卷积一个像素点的感受野，feat2中一个点对应feat1中3×3区域的信息，则生成的feat2的感受野为7×7；第三次处理同上，第二次空洞卷积的卷积核尺寸等于第三次空洞卷积一个像素点的感受野，生成的feat3中每个点对应feat1中7×7的信息，采用空洞率为3的卷积，feat3每一点的感受野为15×15。空洞率对应卷积核尺寸计算如公式（3）所示。

$$K_{new} = k_{ori} + (k_{ori} - 1)(rate - 1) \quad (3)$$

其中，K_{new}为新卷积核尺寸；k_{ori}为原始卷积

核尺寸；*rate*为空洞率大小。

ASPP由4个不同空洞率的空洞卷积和一个特征融合层组成。在同一Input Feature Map的基础上，并行使用4个空洞卷积，空洞率设置为$r=\{6, 12, 18, 24\}$，核尺寸为3×3，最后将不同卷积层得到的结果做像素加融合并到一起。

ASPP+在ASPP的基础上向每个空洞卷积后面添加了BN层和把ASPP中的普通卷积替换为深度可分离卷积（Deep Separable Convolution，DSC），ASPP+和ASPP空洞率也有所不同。主要包含以下几部分：①一个1×1卷积层，以及三个3×3的空洞卷积，其rate设置为｛6，12，18｝，并且含有BN层；②一个全局平均池化得到image-level特征，然后输入1×1卷积，再采用双线性差值还原到原始大小；③将①和②得到的4个不同尺度的特征图在通道维度合并在一起，然后输入1×1的卷积进行融合得到256通道的新特征图。

2.4 评价指标

2.4.1 果园提取效果评价指标

使用像素精度（Pixel Accuracy，PA）、频权交并比（Frequency Weighted Intersection over Union，FWIoU）、均交并比（Mean Intersection over Union，MIoU）、均类像素精度（Mean Pixel Accuracy，MPA）作为提取效果评价指标。苹果园提取被视为语义分割问题，将苹果园像素取0，背景像素取1。

苹果园提取的F_1值如公式（4）所示。

$$F_1 = 2 \times \frac{P \times R}{P + R} \times 100\% \tag{4}$$

其中，P、R分别为苹果园该类别提取的精确率和召回率。

像素精度计算如公式（5）所示。

$$\mathrm{PA} = \frac{\sum_{i=0}^{k} p_{ij}}{\sum_{i=0}^{k}\sum_{j=0}^{k} p_{ij}} \times 100\% \tag{5}$$

频权交并比计算如公式（6）所示。

$$FWIoU = \frac{1}{\sum_{i=0}^{k}\sum_{j=0}^{k} p_{ij}} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}} \times 100\% \tag{6}$$

均交并比计算如公式（7）所示。

$$\mathrm{MIoU} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}} \times 100\% \tag{7}$$

均类像素精度计算如公式（8）所示。

$$\mathrm{MAP} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ij}}{\sum_{j=0}^{k} p_{ij}} \times 100\% \tag{8}$$

其中，k表示目标的类别数，共有$k+1$个类别（含目标和背景）；i、j均表示类别号；p_{ii}为正确分类的像素；p_{ij}和p_{ji}均表示分类错误的像素。

2.4.2 面积精度评价

选取长武县王东沟和白水县通积村这两个研究区对语义分割模型提取的苹果园进行面积精度评价。目视解译出长武县王东沟和白水县通积村苹果园面积分别为139.41和44.97 hm^2，因此将模型提取苹果园面积与目视解译面积进行比较分析。本研究提出一种基于Python第三方开源库PIL和OpenCV的方法对语义分割方法提取的苹果园进行面积计算。首先使用PIL和OpenCV将RGB图像转换为灰度图像，再利用numpy.where的广播机制统计“苹果园”这一类别的像元数，进而计算语义分割方法提取的苹果园面积，最后计算其与目视解译面积的相对精度，并进行分析。其计算如公式（9）和公式（10）所示。

$$A_i = Pr \times A \tag{9}$$

其中，A_i为某类别面积，hm^2；Pr为该类像元占比（Pixel ratio）；A为该区域总面积，hm^2。

$$Pre = 1 - \frac{|A_1 - A_0|}{A_0} \times 100\% \tag{10}$$

其中，Pre为面积精度；A_1为不同语义分割方法提取的苹果园面积，hm^2；A_0为目视解译苹果园面积，hm^2。

2.5 试验参数设置

硬件平台为NVIDIA GeForce RTX 2080S

（8G）GPU、64GB内存和i7-9700k CPU，本研究使用PyTorch框架搭建语义分割网络，初始学习率为1×10^{-4}，优化器选择Adam，迭代次数设置为30次，为防止训练后期学习率过大导致网络参数错过全局最优解，调整策略为每迭代20个epoch之后，学习率变为原来的0.5倍，权重衰减率设为10^{-4}，动量设为0.9。

3 结果与分析

3.1 R_34_Linknet与多种分割算法对比

为验证R_34_Linknet的有效性，本研究将其与多种语义分割算法，包括Linknet、SegNet、FCN_8s、Unet和DeeplabV3+网络做了对比实验。参数初始化方法采用He_uniform初始化方式，从［−limit，limit］中的均匀分布中抽取样本，其中limit是sqrt（6 / fan_in），fan_in是权值张量中的输入单位的数量。试验中网络的损失函数均采用NLLLoss。

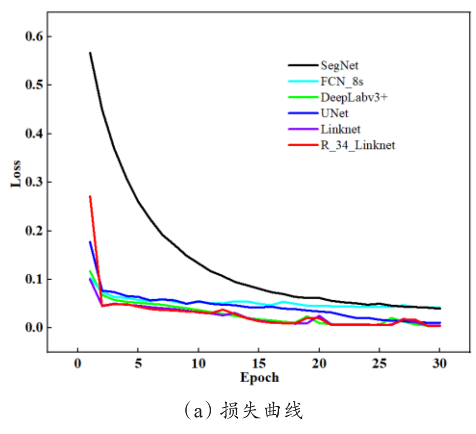
不同模型训练时的损失曲线和MIoU曲线如图2所示，在测试集上表现如表2所示。可知，R_34_Linknet与其他语义分割网络相在训练损失曲线和测试集上各指标表现有一定优势，其损失曲线收敛速度较快，且在经过最后一次迭代时其值最小为0.004；在测试集上MIoU较SegNet、FCN_8s、DeeplabV3+、UNet、原始Linknet网络分别提高13.6%、1.2%、5.4%、7.4%、1.2%，像素精度分别提高6.3%、4.9%、2.4%、3.1%、0.6%。

图3为多种网络在测试集上的分割效果。由图3可知，SegNet误提了许多其他植被、裸地、田间道路等，在一些窄小的地块容易出现漏提，提取结果较为杂乱；DeeplabV3+和UNet相对SegNet有较大提升，但同样误提了裸地、其他植被等。Linknet的苹果园提取效果相对较好，减少了裸地、其他植被的干扰，但在地块边界细节处理上较差，同时也没能完全避免田间道路的影响。R_34_Linknet相比于其它网络，提取的苹果园受到田间道路、其他植被影响较小，在田块边界细节上有了较大提升。SegNet、FCN_8s、DeeplabV3+、UNet、原始Linknet网络和R_34_Linknet在苹果园这一类别的提取精度分别为86.3%、87.2%、89.5%、88.7%、92.0%和92.8%，在苹果园这一类别提取精度R_34_Linknet高于其它网络。

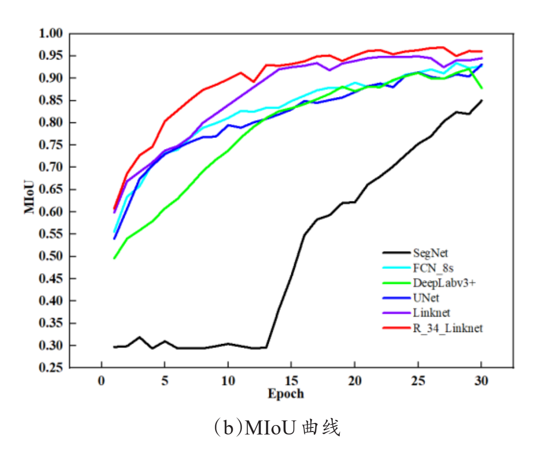
表2 各模型在测试集上的表现

Table 2 Performances of models on test set

分割方法	F_1/%	PA/%	MIoU/%	FWIoU/%	mPA/%
SegNet	84.2	86.0	67.6	76.0	77.9
FCN_8s	85.0	87.4	69.2	76.9	78.6
DeeplabV3+	85.9	89.9	75.8	81.2	88.3
UNet	84.6	89.2	73.8	80.7	82.4
Linknet	86.3	91.7	80.0	85.0	88.6
R_34_linknet	87.1	92.3	81.2	85.7	89.6



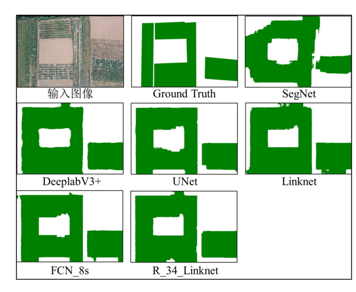
（a）损失曲线



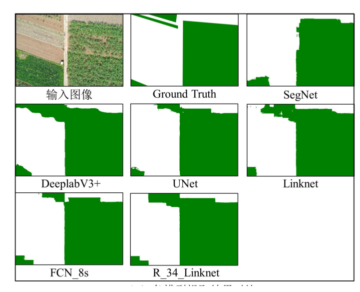
（b）MIoU曲线

图2 不同模型训练损失曲线和MIoU曲线

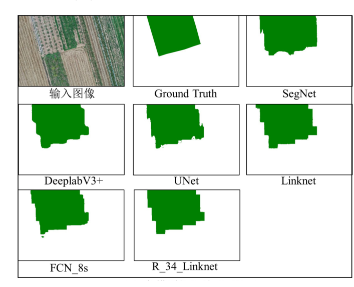
Fig. 2 Training loss curves and MIoU curves of different models



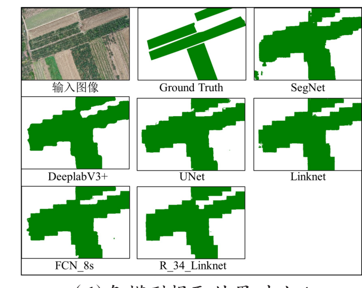
(a)各模型提取结果对比1



(b)各模型提取结果对比2



(c)各模型提取结果对比3



(d)各模型提取结果对比4

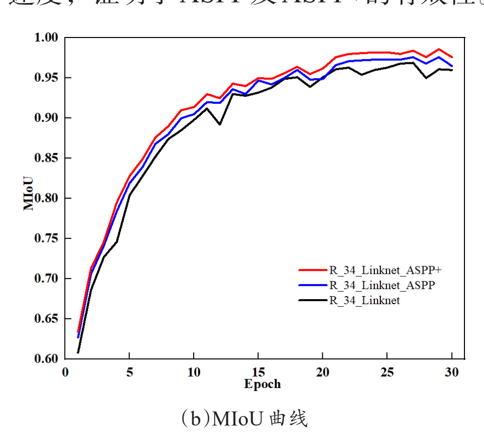
图3 不同输入图像各模型提取结果对比

Fig.3 Comparison of extraction results of the models with different import pictures

3.2 ASPP 和 ASPP+ 加入网络 R_34_Linknet前后对比

为验证ASPP和ASPP+在无人机遥感影像上苹果园提取的效果，本研究将二者分别加入R_34_Linknet 得到 R_34_Linknet_ASPP 和 R_34_Linknet_ASPP+网络，并对比研究了加入前后的提取精度。其训练损失曲线和MIoU曲线变化如图4所示，各项评价指标如表3所示。加入ASPP之后MIoU提高2.1%，像素精度提高1.1%，加入ASPP+之后MIoU和像素精度又分别提高2.2%和1.3%。由图4可知，ASPP能够有效提高网络的MIoU，同时没有影响网络训练时的收敛速度，证明了ASPP及ASPP+的有效性。

(a)损失曲线



(b)MIoU曲线

图4 三种改进模型训练损失曲线和MIoU曲线

Fig. 4 Training loss curves and MIoU curves of three improved models

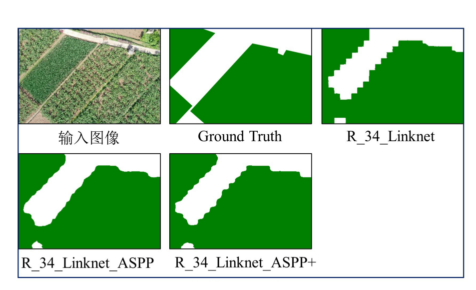
通过训练损失曲线图2和图4可以看出，各模型的损失值随着训练次数的增加而减小，MIoU值随着训练次数的增加而提高。由表2和表3可知本研究改进算法R_34_Linknet_ASPP+在测试集上各指标表现最好。图5为各网络在测试集上的提取效果，可以看出，图5（a）、图5（b）和图5（c）的R_34_Linknet_ASPP和R_34_Linknet_ASPP+较R_34_Linknet提取的苹果园

表3 三种模型各项评价指标

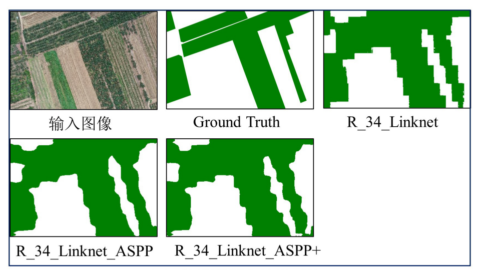
Table 3 Evaluation indicators of three models

分割方法	F_1/%	PA/%	MIoU/%	FWIoU/%	MPA/%
R_34_Linknet	87.1	92.3	81.2	85.7	89.6
R_34_Linknet_ASPP	87.9	93.4	83.3	87.7	90.6
R_34_Linknet_ASPP+	88.3	94.7	85.5	89.0	92.3

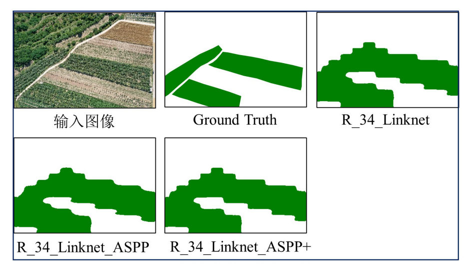
地块边缘处更加平滑，准确。通过计算图5（d）R_34_Linknet_ASPP 和 R_34_Linknet_ASPP+ 提取出的苹果园像元总数比R_34_Linknet提取出的苹果园像元总数更接近Ground Truth中苹果园像元总数。



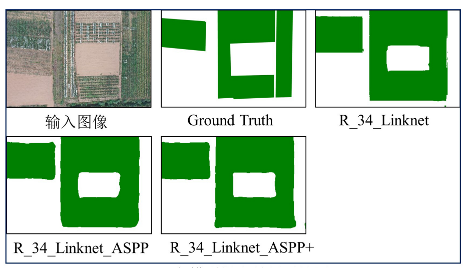
(a)各模型提取结果对比1



(b)各模型提取结果对比2



(c)各模型提取结果对比3



(d)各模型提取结果对比4

图5 各模型提取结果对比

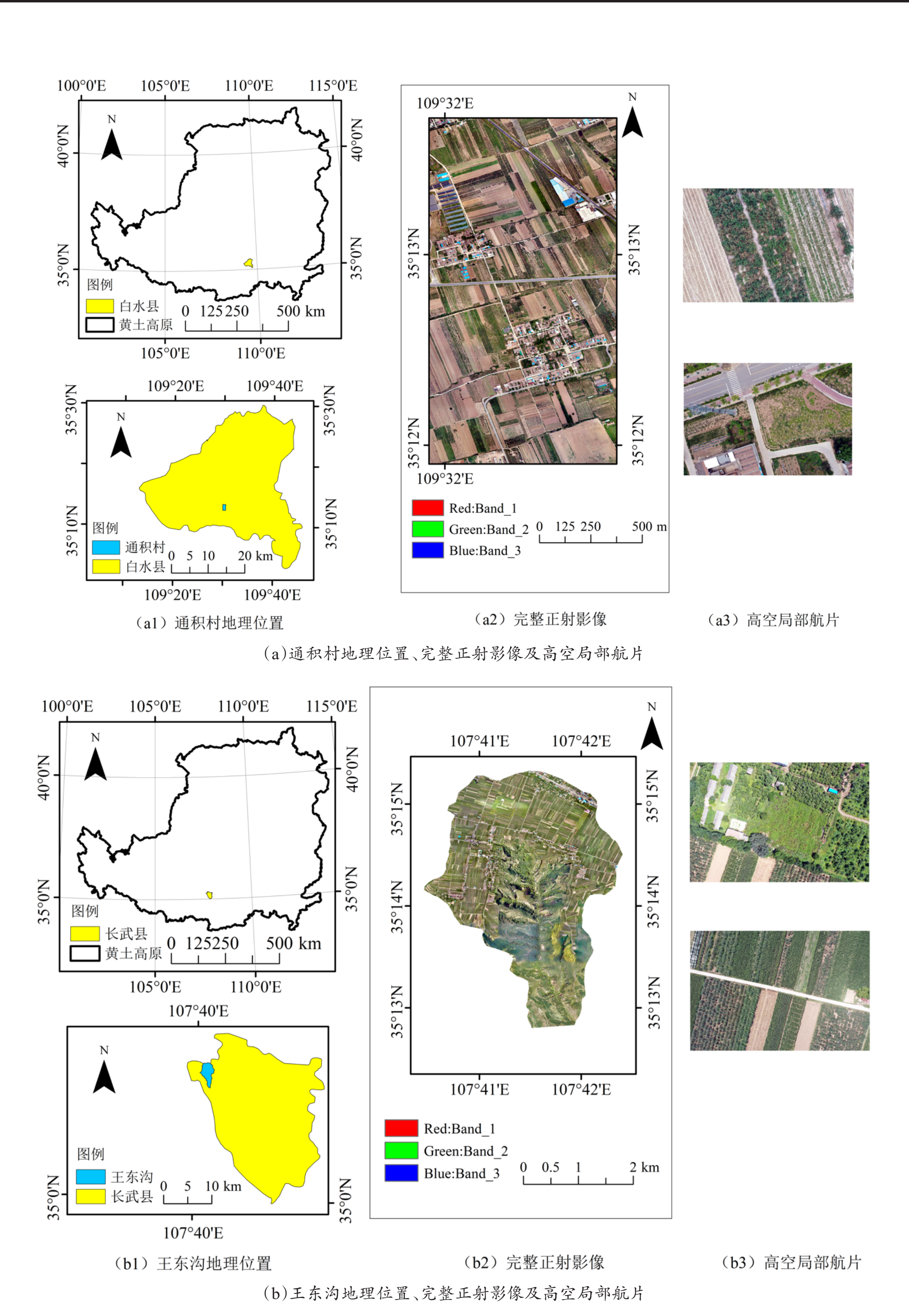
Fig. 5 Extraction results comparison of the models

本研究改进的深度学习语义分割算法R_34_Linknet_ASPP+，在各种复杂背景下，有很好的提取效果。其分割结果对比如图3和图5所示，分析发现其精度较高的原因是把Rest-Net34 [21] 迁移到Linknet网络并在编码器和解码器之间添加了ASPP+ [24] 结构，如图5中R_34_Linknet_ASPP+对图像中边界提取的准确性最好，说明随着训练的进行，该模型学习到了关于苹果园较多有效的特征，在预测时边界处的像素点能够较好地紧密相连。如图5（a）中输入图像含有抽穗期的小麦且颜色纹理与周围苹果园较为相似；图5（b）和图5（c）中两张输入图像含有成熟期的夏玉米、收割之后的麦茬地及裸地，且这三种土地类型较为相似；图6（d）输入图像中的苹果园里覆有反光膜且较为明亮。模型通过对复杂背景中标签图像的学习，更加适用于实际环境，也具有更强的鲁棒性。

3.3 R_34_Linknet_ASPP+的应用

经过3.1和3.2的分析可知，R_34_Linknet_ASPP+相较于其它分割算法各指标表现最优，为进一步验证R_34_Linknet_ASPP+的实际分割效果，利用该模型对长武县王东沟小流域和白水县通积村这两个区域苹果园的空间分布进行提取，并对提取面积进行精度分析。

使用PhotoScan1.4.5软件对长武县王东沟和白水县通积村的航片进行拼接 [32, 33]。长武县王东沟和白水县通积村的地理位置、完整正射影像以及高空局部航片如图6所示，其完整正射影像面积分别为8.3和1.5 km^2，本节所用数据采集时间分别为2021年6月21日至22日（长武县王东沟）、2021年6月23日（白水县通积村）。



（a1）通积村地理位置　（a2）完整正射影像　（a3）高空局部航片

(a)通积村地理位置、完整正射影像及高空局部航片

（b1）王东沟地理位置　（b2）完整正射影像　（b3）高空局部航片

(b)王东沟地理位置、完整正射影像及高空局部航片

注:基于自然资源部标准地图服务网站GS(2016)1550号标准地图制作,底图边界无修改

图6　通积村和王东沟地理位置、完整正射影像及高空局部航片

Fig. 6 Geographical locations, complete orthophotos and high-altitude partial aerial photographs of Wangdonggou and Tongji Village

各模型对长武县王东沟和白水县通积村苹果园面积提取精度如表4和表5所示。

表4　长武县王东沟苹果园面积提取精度

Table 4 Area extraction accuracy of Wangdonggou apple orchard in Changwu County

分割方法	提取面积/hm²	目视解译面积/hm²	面积精度/%
SegNet	157.17		87.26
FCN_8s	125.48		90.01
DeeplabV3+	127.30		91.32
UNet	150.48	139.41	92.06
Linknet	129.21		92.68
R_34_linknet	129.67		93.01
R_34_Linknet_ASPP	130.54		93.64
R_34_Linknet_ASPP+	131.35		94.22

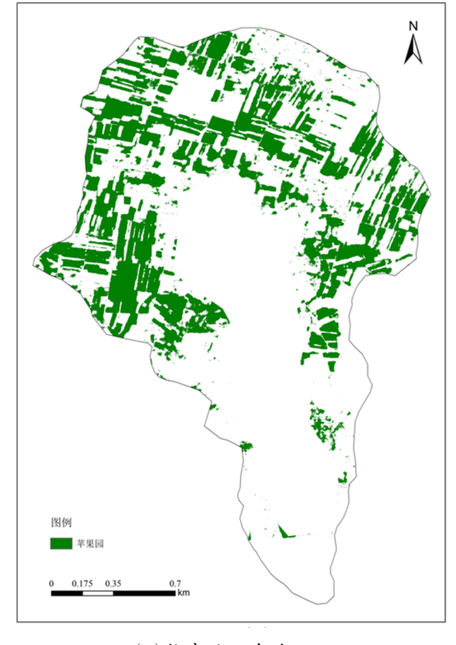
从表4和表5中结果可以看出，同一模型在长武县王东沟和白水县通积村苹果园面积提取精度方面表现基本相同，是由于两地的种植结构以及正射影像中的地物类型基本相同。在王东沟和通积村都表现出R_34_Linknet、R_34_Linknet_ASPP、R_34_Linknet_ASPP+这三种语义分割方法整体上比另外五种语义分割方法精度高。本研究三种改进算法中，R_34_Linknet_ASPP+精度最高，在王东沟和通积村苹果园面积提取精度分别为94.22%和95.46%，在王东沟分别比R_34_Linknet、R_34_Linknet_ASPP高出1.21%和0.58%；在通积村分别比R_34_Linknet、R_34_Linknet_ASPP高出1.70%和0.90%。

表5　白水县通积村苹果园面积提取精度

Table 5 Aera extraction accuracy of apple orchard area in Tongji Village, Baishui County

分割方法	提取面积/hm²	目视解译面积/hm²	面积精度/%
SegNet	49.62		89.65
FCN_8s	41.04		91.27
DeeplabV3+	41.94		93.26
UNet	48.31	44.97	92.57
Linknet	42.16		93.76
R_34_linknet	42.24		93.94
R_34_Linknet_ASPP	42.52		94.56
R_34_Linknet_ASPP+	42.93		95.46

基于R_34_Linknet_ASPP+语义分割方法提取的长武县王东沟和白水县通积村苹果园空间分布如图7所示。



(a)长武县王东沟



(b)白水县通积村

图7　苹果园空间分布提取结果

Fig. 7 Extraction results of the spatial distribution of apple orchards

4 结　论

本研究针对黄土高原苹果园分布信息提取，提出了一种更加高效精准的提取方法，保证了苹果园提取的准确性，提取到的苹果园地块边缘处效果更好。

（1）根据黄土高原无人机影像的特点，制作了面向该区无人机低空遥感影像苹果园识别的专业数据集，包含各种场景的苹果园无人机遥感影像。

（2）融合迁移学习与深度学习方法，将残差神经网络ResNet34网络迁移到Linknet网络，得到R_34_Linknet网络。将R_34_Linknet网络与5种常用的深度学习语义分割模型SegNet、FCN_8s、DeeplabV3+、UNet和Linknet应用于黄土高原苹果园空间分布提取，表现最好的模型为R_34_Linknet，其在测试集上的调和平均值F_1为87.1%，像素准确度PA为92.3%，均交并比MIoU为81.2%，频权交并比FWIoU为85.7%，平均像素准确度MPA为89.6%。

（3）将空间金字塔池化结构（Atrous Spatial Pyramid Pooling，ASPP）与R_34_Linknet网络相结合，扩大网络的感受野，得到R_34_Linknet_ASPP网络；然后对ASPP结构进行改进，得到R_34_Linknet_ASPP+网络。对比三种网络性能，表现最优的为R_34_Linknet_ASPP+，在测试集上F_1为86.3%，PA为94.7%，MIoU为82.7%，FWIoU为89.0%，MPA为92.3%。使用R_34_Linknet_ASPP+在长武县王东沟和白水县通积村提取苹果园面积精度分别为94.22%和95.66%。在王东沟分别比R_34_Linknet、R_34_Linknet_ASPP高出1.21%和0.58%；在通积村分别比R_34_Linknet、R_34_Linknet_ASPP高出1.70%和0.90%。

参考文献：

[1] 周江涛, 赵德英, 陈艳辉, 等. 中国苹果产区变动分析[J]. 果树学报, 2021, 38(3): 372-384.
ZHOU J, ZHAO D, CHEN Y, et al. Analysis of apple producing area changes in China[J]. Journal of Fruit Science, 2021, 38(3): 372-384.

[2] GAO X, ZHAO X, WU P, et al. The economic-environmental trade-off of growing apple trees in the drylands of China: A conceptual framework for sustainable intensification[J]. Journal of Cleaner Production, 2021 (3): 126-497.

[3] 康凌艳, 雷玉平, 郑力, 等. 在GIS支持下利用MODIS数据监测多种作物和果树种植面积[J]. 遥感技术与应用, 2007(3): 361-366.
KNAG L, LEI Y, ZHENG L, alel. Vegetation classification based on modis data and the accuracy evaluation in the pixel scale[J]. Remote Sensing Technology and Application, 2007(3): 361-366.

[4] 邬明权, 杨良闯, 于博, 等. 基于遥感与多变量概率抽样调查的作物种植面积测量[J]. 农业工程学报, 2014, 30(2): 146-152.
WU M, YANG L, YU B, et al. Mapping crops acreages based on remote sensing and sampling investigation by multivariate probability proportional to size[J]. Transactions of the CSAE, 2014, 30(2): 146-152.

[5] 刘佳岐. 基于Landsat8遥感影像的扶风县苹果园地信息提取研究[D]. 杨凌: 西北农林科技大学, 2015.
LIU J. Research on apple oechards information extraction of fufeng countybased on landsat8 remote sensing image[D]. Yangling: Northwest A&F University, 2015.

[6] KUMAR A, SINGH K, LAL B, et al. Mapping of apple orchards using remote sensing techniques in cold desert of himachal pradesh, India[J]. Journal of the Indian Society of Remote Sensing, 2008, 36(4): 387-392.

[7] YUAN H, MA R, LUO J. Mapping orchards on plain terrains using multi-temporal medium-resolution satellite imagery[J]. Applied Engineering in Agriculture, 2015, 31: 351-362.

[8] 徐晗泽宇, 刘冲, 王军邦, 等. Google Earth Engine平台支持下的赣南柑橘果园遥感提取研究[J]. 地球信息科学学报, 2018, 20(3): 396-404.
XU H, LIU C, WANG J, et al. Study on extraction of citrus orchard in Gannan region based on Google Earth engine platform[J]. Journal of Geo-information Science, 2018, 20(3): 396-404.

[9] 万亚玲, 钟锡武, 刘慧, 等. 卷积神经网络在高光谱图像分类中的应用综述[J]. 计算机工程与应用, 2021. 57(4): 1-10.
WAN Y, ZHONG X, LIU H, et al. Survey of application of convolutional neural network in classification of hyperspectral images[J]. Computer Engineering and Applications, 2021, 57(4): 1-10.

[10] BERA S, SHRIVASTAVA V K. Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification[J]. International Journal of Remote Sensing, 2020, 41: 2664-2683.

[11] MAJEED Y, ZHANG J, ZHANG X, et al. Deep learning based segmentation for automated training of apple trees on trellis wires[J]. Computers and Electronics in Agriculture, 2020, 170: ID 170.

[12] JEONGEUN K, JEAHWI S, SUKWOO L, et al. An intelligent spraying system with deep learning-based semantic segmentation of fruit trees in orchards[C]// IEEE International Conference on Robotics and Automation (ICRA). Piscataway, New York, USA: IEEE, 2020: 3923-3929.

[13] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39: 640-651.

[14] OLAF R, PHILIPP F, THOMAS B. U-Net: convolutional networks for biomedical image segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Piscataway, New York, USA: IEEE, 2015: 234-241.

[15] 马金林, 邓媛媛, 马自萍. 肝脏肿瘤CT图像深度学习分割方法综述[J]. 中国图象图形学报, 2020, 25(10): 2024-2046.

MA J, DENG Y, MA Z. Review of deep learning segmentation methods for CT images of liver tumors[J]. Journal of Image and Graphics, 2020, 25(10): 2024-2046.

[16] 张凯航, 冀杰, 蒋骆, 等. 基于SegNet的非结构道路可行驶区域语义分割[J]. 重庆大学学报, 2020, 43(3): 79-87.

ZHANG K, JI J, JIANG L, et al. The segmentation of driving regions on unstructured road based on SegNet architecture[J]. Journal of Chongqing University, 2020, 43(3): 79-87.

[17] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J/OL]. arXiv: 1409.1556, 2014.

[18] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[J/OL]. arXiv: 1603. 05027, 2016.

[19] ANAGNOSTIS A, TAGARAKIS A C, ASIMINARI G, et al. A deep learning approach for anthracnose infected trees classification in walnut orchards[J]. Computers and Electronics in Agriculture, 2021, 182: ID 182.

[20] PAPANDREOU G, CHEN L, KEVIN P, et al. Weakly- and semi-supervised learning of a deep convolutional network for semantic image segmentation[C]// IEEE International Conference on Computer Vision. Piscataway, New York, USA: IEEE, 2015: 1742-1750.

[21] CHEN L, PAPANDREOU G, KOKKINOS I. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 40: 834-848.

[22] CHEN L, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J/OL]. arXiv:1706.05587v3, 2017.

[23] CHAURASIA A, CULURCIELLO E. LinkNet: exploiting encoder representations for efficient semantic segmentation[C]// IEEE Visual Communications and Image Processing. Piscataway, New York, USA: IEEE, 2017.

[24] DENG W, MOU Y, KASHIWA T, et al. Vision based pixel-level bridge structural damage detection using a link ASPP network[J]. Automation in Construction, 2020, 110: ID 102973.

[25] YU Y, LI H, WANG J, et al. A multilayer pyramid network based on learning for vehicle logo recognition[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(5): 3123-3134.

[26] MO J, LAN Y, YANG D, et al. Deep learning-based instance segmentation method of litchi canopy from UAV-acquired images[J]. Remote Sensing, 2021, 13 (19): ID 3919.

[27] 王爱. 黄土高原苹果树识别与蒸散发过程模拟[D]. 杨凌: 西北农林科技大学, 2020.

WANG A. Apple tree identification and evapotranspiration process simulation on the Loess Plateau[D]. Yang ling: Northwest A&F University, 2020.

[28] 王培周. 黄土高原苹果主产区养分投入和土壤养分状况及其空间分布特征[D]. 杨凌: 西北农林科技大学, 2021.

WANG P. Nutrient input, soil nutrient contents and their spatial distribution characteristics in the main apple producing areas of the Loess Plateau[D]. Yangling: Northwest A&F University, 2021.

[29] 晏利斌. 1961—2014年黄土高原气温和降水变化趋势[J]. 地球环境学报, 2015, (5): 276-282.

YAN L. Characteristics of temperature and precipitation on the Loess Plateau from 1961 to 2014[J]. Journal of Earth Environment, 2015, (5): 276-282.

[30] 吴乾慧, 张勃, 马彬, 等. 气候变暖对黄土高原冬小麦种植区的影响[J]. 生态环境学报, 2017, 26: 429-436.
WU Q, ZHANG B, MA B, et al. Impact of climate warming on winter wheat planting in the Loess Plateau[J]. Ecology and Environmental Sciences, 2017, 26: 429-436.
[31] CHEN L, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[J]. Computer vision-ECCV, 2018, 11211: 833-851.
[32] 陈鹏飞, 徐新刚. 无人机影像拼接软件在农业中应用的比较研究[J]. 作物学报, 2020, 46(7): 1112-1119.
CHEN P, XU X. A comparison of photogrammetric software packages for mosaicking unmanned aerial vehicle (UAV) images in agricultural application[J]. Acta Agronomica Sinica, 2020, 46(7): 1112-1119.
[33] 王慧. PhotoScan在无人机遥感影像数据处理中的应用[J]. 测绘与空间地理信息, 2017, 40(5): 109-111.
WANG H. Application of PhotoScan in UAV remote sensing image data processing[J]. Geomatics & Spatial Information Technology, 2017, 40(5): 109-111.

Accurate Extraction of Apple Orchard on the Loess Plateau Based on Improved Linknet Network

ZHANG Zhibo[1], ZHAO Xining[2*], GAO Xiaodong[2], ZHANG Li[1], YANG Menghao[2]

(1. *College of Water Resources and Architectural Engineering*, *Northwest A & F University*, *Yangling* 712100, *China*;
2. *Institute of Soil and Water Conservation*, *CAS & MWR*, *Yangling* 712100, *China)*

Abstract: The rapid increasing of apple planting area on the Loess Plateau has exerted an important influence on the regional eco-hydrology and socio-economic development. However, the orchards in this area are small and complex, and there are only county or city scale statistical data, lack of actual spatial distribution information. To this end, for the extraction of apple orchards on the Loess Plateau, in this study, a professional dataset of low-altitude remote sensing images acquired by unmanned aerial vehicle was firstly established. The R_34_Linknet network and other five commonly used deep learning semantic segmentation models SegNet, FCN_8s, DeeplabV3+, UNet and Linknet were applied to the spatial distribution extraction of apple orchards on the Loess Plateau, and the best-performing model was R_34_Linknet, with a F1 score of 87.1%, a pixel accuracy (PA) of 92.3%, an mean intersection over union (MioU) of 81.2%, a frequency weighted intersection over union (FWIoU) of 85.7%, and the mean pixel accuracy (MPA) was 89.6%. The spatial pyramid pool structure (ASPP) and R_34_Linknet network was combined to expand the receptive field of the network and get R_34_Linknet_ASPP network, and then ASPP structure was improved. Combining the spatial pyramid pooling (ASPP) with the R_34_Linknet network to expand the receptive field of the network and obtain a R_34_Linknet_ASPP network; Then the ASPP structure was improved to get a R_34_Linknet_ASPP+ network. The performance of the three networks were compared. R_34_Linknet_ASPP+ got the best performance, with 86.3% for F_1, 94.7% for PA, 82.7% for MIoU, 89.0% for FWIoU, and 92.3% for MPA on the test set. The accuracy of apple orchard extraction in Wangdonggou, Changwu County and Tongji Village, Baishui County using R_34_Linknet_ASPP+ were 94.22% and 95.66%, respectively. In Wangdonggou, it was 1.21% and 0.58% higher than R_34_Linknet and R_34_Linknet_ASPP, respectively. In Tongji village, it was 1.70% and 0.90% higher than R_34_Linknet and R_34_Linknet_ASPP, respectively. The results show that the proposed R_34_Linknet_ASPP+ method can extract apple orchards accurately, the edge treatment of apple orchard plots is better, the method can be used as the technical support and theoretical basis for research on the spatial distribution mapping of apple orchards on the Loess Plateau.

Key words: UAV remote sensing; apple orchard extraction; deep learning; Loess Plateau; transfer learning; residual neural network; semantic segmentation